EJFA Emirates Journal of Food and Agriculture

Emirates Journal of Food and Agriculture 37: 1–12
doi: 10.3897/ejfa.2025.153356

RESEARCH PAPER FEEDING THE FUTURE

Rapid prediction of green tea quality based on convolutional neural network*

Danyi Huang[1,2], Lamei Li[3], Yimin Xuan[2], Dongmei Fan[4], Xiaochang Wang[2,5]

1 *College of Cooperative Economics, Zhejiang Institute of Economics and Trade, Hangzhou 310018, China*
2 *Tea Research Institute, Zhejiang University, Hangzhou 310058, China*
3 *Shaoxing Economic Crop Technology Promotion Center, Shaoxing, 312030, China*
4 *College of Tea Science and Tea Culture, Zhejiang A and F University, Hangzhou 311300, China*
5 *Institute of China Dafo Longjing Tea, Shaoxing 312599, China*

Corresponding author: Xiaochang Wang (xcwang@zju.edu.cn)

Academic editor: Zeynep Ünal ♦ **Received** 19 March 2025 ♦ **Accepted** 18 June 2025 ♦ **Published** 30 September 2025

Abstract

The appearance of tea is a core visual indicator of quality evaluation, which directly affects consumer perception and market grading. This study proposed a green tea appearance quality evaluation method based on a convolutional neural network. A green tea image dataset of three shapes was constructed, containing flat-shaped Dafo Longjing tea, sparrow tongue-shaped Huangshan Maofeng tea, and granular-shaped Pingshui Rizhu tea. Dafo Longjing tea and Huangshan Maofeng tea had six grades and Pingshui Rizhu tea had four grades. Approximately 1500 images were collected for each grade of tea samples, and 9608, 9312, and 5768 images were obtained for three kinds of teas, respectively. The AdamX optimizer, ExponentialLR learning rate decay and batch size 32 were used to compare the classification performance of the ResNet18 model and MobileNetV3 model. The results showed that ResNet18 model achieved 98.28%, 99.19%, and 99.74% accuracies in recognizing grade in Dafo Longjing, Huangshan Maofeng, and Pingshui Rizhu. And MobileNetV3 model achieved 98.91%, 99.62%, and 99.39%, respectively. The fluctuation trends of the precision and recall in the two models during grade identification were consistent, reaching more than 98.20% in ResNet18 model and more than 98.90% in MobileNetV3 model. The effect of parameter tuning on performance metrics was similar in three green teas, indicating that these two models were appropriate for usage in various shapes of green tea. The ResNet18 model with the ECA attention mechanism was larger than the MobileNetV3 model, so MobileNetV3 is more suitable for actual quality detection with its lightweight advantage. Both models confirm the effectiveness of deep learning in tea appearance quality evaluation and provide technical support for automated grading.

Keywords

green tea, appearance, image recognition, grade, ResNet18, MobileNetV3

Introduction

As a non-fermented tea, green tea is highly favored for its unique flavor profile, rich bioactive components (e.g., tea polyphenols, catechins), and significant health benefits (Madhura et al. 2016; Yang et al. 2020; Choi et al. 2023; Connolly et al. 2024). In the traditional Chinese tea-making process, green tea has diversified appearance characteristics through the key processes of fixation, rolling, and drying, such as the flat form of Longjing tea, the granular structure of Gunpowder tea, and the needle shape of Yulu tea. These morphological differences not only reflect the tenderness of the raw materials and the level of processing technology but are also directly related to the sensory quality of the tea

* Topical Collection "Feeding the Future: Advancements in Deep Learning for Sustainable Agriculture and Food Security", edited by Prof. Yonis Gulzar and Zeynep Ünal.

Copyright *Huang, et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and the market value of the grade. However, there are many different types and grades of green tea, which adds to the difficulty of determining quality. Screening of key components in tea helps to determine quality (Li et al. 2023; Santos et al. 2023; Yang et al. 2023; Lu et al. 2024), but the procedure is complex and time-consuming to perform. It is easier to evaluate tea based on appearance. The use of spectroscopic techniques for tea categorization and quality identification is well established. Green tea origins, kinds, and grades can be effectively classified using multi-spectral imaging techniques (Wu et al. 2008), hyperspectral imaging techniques, near infrared imaging techniques, 1H Nuclear Magnetic Resonance (NMR) (Hou et al. 2024) and so on. An RBF model based on RGB color patterns was developed to predict the sensory evaluation (Wu et al. 2008). The use of hyperspectral imaging technology in matcha processing allowed pigment visualization and was useful in tea quality determination (He et al. 2024). Green tea grades can be accurately identified using multisensor information, such as spectral, visual, and olfactory features (Li et al. 2018). Currently, tea quality evaluation faces two major challenges: first, relying on human sensory evaluation is susceptible to subjective factors; second, although modern instrumental analytical techniques such as spectroscopy and chromatography can accurately identify components, there are limitations such as high equipment costs, complex operations, and difficulties in real-time online detection. Intelligent recognition technology based on appearance morphology has become an important direction to break through the bottleneck of traditional inspection. As a comprehensive characterization of tenderness, process integrity, and quality grade, tea appearance has the advantages of intuitive, non-destructive, and quantifiable assessment. Recently, there has been a tremendous advancement in image identification technology, moving from manual feature extraction-based methods to deep learning approaches, with convolutional neural networks (CNN) demonstrating strong performance (Krizhevsky et al. 2012). It has fewer requirements for devices that can adaptively choose the derived recognition features throughout the iterative training process. Deep learning was widely used in agriculture, especially for the classification of plants (Amri et al. 2024; Chouhan et al. 2024; Gulzar et al. 2024) and foods (Zhou et al. 2019; Gulzar 2024; Gulzar and Ünal 2025b) with good performance. Image recognition technologies combined with deep learning have been used in the management of tea gardens and the evaluation of tea quality. For example, identifying pests and pathogens in fresh tea leaves (Chen et al. 2019), forecasting tea yield (Wang et al. 2024a), and accurately selecting and identifying images of fresh tea leaves and buds (Wang et al. 2024b; Wu et al. 2024). Very high resolution (VHR) remote sensing images can extract images of tea gardens to create maps (Tang et al. 2020). The identification of the Longjing tea grade was accomplished by a camera picture based on deep CNN and transfer learning (Zhang et al. 2023). The use of image enhancement may create the black tea shape quality recognition accuracy of 94.5%. Jiang et al. proposed an artificial intelligence method for tea grade classification based on a symmetry detection algorithm (Jiang and Chen 2021). Three-grade orthodox black tea can be classified with 97.22% accuracy using an Improved Inception Network (Guo et al. 2024). 83 image information was extracted from black and green tea, and decision tree models with recognition accuracy of 92.92% and 95.00% were obtained (Bakhshipour et al. 2020).

There is still room for improvement in the current technology system: (1) most of the studies focus on a single tea type or a specific processing form, and lack generalized modeling of the appearance characteristics of multiple green tea types; (2) the traditional image processing methods rely on artificial feature engineering, which makes it difficult to capture complex textures and spatial correlations; and (3) the variations in lighting, background noise, and gesture diversity of the actual production scene lead to insufficient model robustness. Responding to the above issues, this study proposed a grade recognition method combining computer vision and deep learning. Tea samples with different shapes were collected in different environments. The appearance features were automatically analyzed through deep learning. The ResNet18 model and MobileNetV3 model proposed in this study can effectively realize the grade recognition of green tea with different shapes. In addition, parameter adjustment, such as selecting AdamX optimizer, ExponentialLR learning rate decay, and batch size 32 helps to improve the performance metrics such as accuracy, precision, recall, and F1-score, without the problem of overfitting. These model optimizations can also support non-experts in making quick and reliable grade assessments of green tea.

- A detailed review has been conducted, examining the most recent works in said area using image recognition and convolutional neural network.
- A green tea image database was constructed by collecting green tea of various appearances and grades in different environments in order to improve the diversity of tea images and the generalizability of validation.
- The problem of grade identification of green tea was reintroduced, and the ResNet18 model and the MobileNetV3 model were selected for grade identification of green tea images.
- Various model tuning techniques were incorporated to improve the performance of the proposed model and to ensure that the model does not overfit.

Materials and methods

Sample preparation

Three different shapes of green tea were selected for this experiment (Fig. 1): Dafo Longjing tea (DFLJ), Huangshan Maofeng tea (HSMF), and Pingshui Rizhu tea (PSRZ). DFLJ was from Xinchang, Zhejiang province. Its buds and leaves were flat-shaped. It was assigned to Special grade (S), First grade (1), Second grade (2), Third grade (3), Fourth grade (4), and Fifth grade (5). HSMF

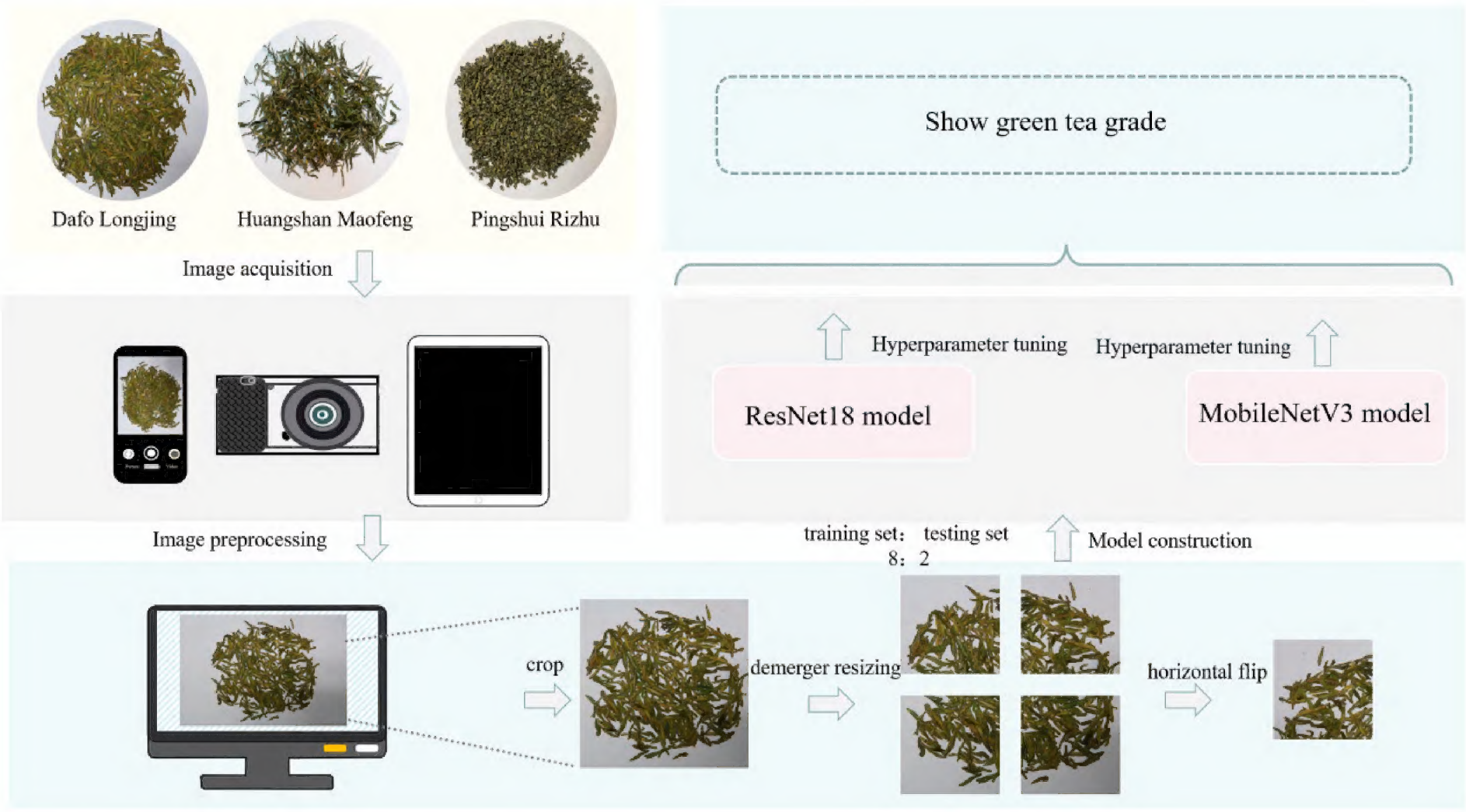

Figure 1. The technical flow chart of this experiment.

was from Huangshan, Anhui province. It had a slightly curled, sparrow tongue shape. It can be categorized into six grades, which were Special first grade (S1), Special second grade (S2), Special third grade (S3), First grade (1), Second grade (2), and Third grade (3). PSRZ Tea was from Shaoxing, Zhejiang province. Tea leaves were rolled into grains. It had four grades as First grade (1), Second grade (2), Third grade (3), Fourth grade (4), respectively.

Tea image acquisition and preprocessing

Tea image acquisition

In order to reduce the impact of various factors, such as equipment, light source, shooting angle, the amount of tea, and other conditions, on the accuracy of the model, five kinds of equipment were used to collect images under multiple scenes. Xiaomi 10Pro and Huawei Enjoy 9S smartphone, Apple iPad Air5, Sony α6100 camera with Sigma 30mm f1.4 lens, and Sony Black Card 5a camera were photo acquisition equipment. Warm light sources and daylight bulbs of different intensities were used as photographic lighting resources. Image acquisition from different heights downward and at an oblique angle of 45° for varying quantities of tea leaves in order to bring variation to the tea images and to verify generalizations.

Image preprocessing

The tea image was preprocessed using Python-OpenCV by calling cv2.resize. After cropping the image to a square, it was divided into four smaller squares using a 2×2 grid (Fig. 1). Every image was proportionally scaled to a 300×300×3 RGB image before being flipped horizontally. As a result, the green tea image database has been correspondingly expanded by four times the original number, 9608 images from the DFLJ, 9312 images from the HSMF, and 5768 images from the PSRZ were acquired. On the other hand, the use of data augmentation techniques prevented excessive scaling from losing details of the tea appearance and improved the model's ability to recognize the target (Gulzar and Ünal 2025a). The dataset was divided into a training set and a testing set of 8:2, and the model was trained and tested several times using five-fold cross-validation to ensure the stability and reliability of the results. And Table 1 showed the number of images per class of green tea. The program additionally used transforms package of PyTorch to normalize the images during model training. This involved evenly resizing them

Table 1. Number of images per grade of green tea.

Green tea	Grade	Number of images	Training set	Testing set
DFLJ	S	1600	1280	320
	1	1600	1280	320
	2	1600	1280	320
	3	1600	1280	320
	4	1600	1280	320
	5	1608	1287	321
HSMF	S1	1552	1242	310
	S2	1552	1242	310
	S3	1544	1236	308
	1	1552	1242	310
	2	1552	1242	310
	3	1560	1248	312
PSRZ	1	1440	1152	288
	2	1440	1152	288
	3	1440	1152	288
	4	1448	1159	289

to 256 pixels in length and width, followed by a random cutting of the images into 224×224 pixels and a random angle rotation. After that, the picture data was normalized and turned into a PyTorch tensor. By lessening the gradient vanishing issue, normalization can increase the network's capacity for learning and enhance the model's training stability and convergence speed.

Model building and performance evaluation

Model construction

The experiment used two convolutional neural network models, ResNet18 and MobileNetV3. ResNet was introduced in 2015 and is characterized by the introduction of residual concatenation, which is often used for image recognition as it reduces redundant information in the data while maintaining a high level of accuracy (Bagherzadeh et al. 2024). MobileNet can achieve high accuracy in a short training time by using a small amount of data to train the model (Elfatimi et al. 2021). This study set up five types of optimizers, three learning rates, five learning rate decays, five batch sizes, seven regularizations and five attention mechanisms to evaluate the grade recognition effect of three types of green tea.

Model evaluation

Grade categorization of green tea based on appearance quality was essentially a multi-classification problem. Sample occurrences are classified into three or more categories in a multiclassification problem. Owing to the limitations of the classification methodology and model, it is imperative to evaluate the accuracy of the classifier's classification results. Classification evaluation indexes employed in this work included the confusion matrix, accuracy, precision, recall, and F1-score (Liu et al. 2010). These metrics were computed using Equations (1) to (4). A confusion matrix is a visualization tool for supervised learning that summarizes data in a dataset into true and predicted categories in the form of a matrix, with the rows representing real values and the columns representing predicted values. This situation analysis chart summarizing the predictions of a classification model might visually assist us determining which group of samples the model performs well inside. Accuracy, precision, recall, and F1-score to evaluate the performance of different deep learning models for green tea appearance quality recognition effects.

$$\text{Accurary} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Training and detection setting

Models were trained on a Windows 10 Professional system using Python 3.8 and PyTorch 2.1.1+cu121. The hardware consisted of an 11th Gen Intel Core i5-11400F CPU and an NVIDIA GeForce RTX 3060 GPU with 16GB of RAM. Data analysis was performed using Origin 2024 and Python 3.8.

Results

ResNet18 model building and green tea grade recognition

Hyperparameter tuning

Three green tea grade recognition models have been established using the ResNet18 network. The effect of five optimizers SGD, Adagrad, Adadelta, Adam, and AdamX on green tea grade identification was examined (Fig. 2a–f). The training results of the different optimizers in the three kinds of green tea followed the similar pattern, with all achieving convergence shortly after 10 rounds of training. The evaluation metrics demonstrated that models trained with the Adagrad and Adadelta optimizers do not perform well on all grade recognition models. In the DFLJ dataset, the optimizers Adam and AdamX outperformed SGD by 1% on all evaluation metrics. The accuracy of the DFLJ grade recognition model was 93.34% and 93.29% for Adam and AdamX, respectively. The HSMF dataset fared the best using the AdamX optimizer, with grade recognition accuracy of 95.32%. The PSRZ dataset performed best with SGD and AdamX optimizers, with an accuracy of 97.23%. Furthermore, when training on different datasets, the AdamX optimizer had higher accuracy and lower loss values. When combined, the AdamX optimizer outperformed others in recognizing the grades of various green tea shapes.

Using AdamX as the optimizer, a green tea grade recognition model was created, and the learning rate was adjusted to improve the recognition accuracy of the model. Initially, the learning rates were set at 0.01, 0.001, and 0.0001 (Fig. 2g). With delayed convergence speed, low accuracy, and huge loss values, it was clear that the models with learning rates of 0.01 and 0.0001 had poor performance in the three green tea datasets. In the three datasets of DFLJ, HSMF, and PSRZ, the ResNet18 model with a learning rate of 0.001 obtained a higher accuracy rate of 95.06%, 96.88%, and 97.57%, respectively. The further adjusted learning rates were 0.002, 0.0008, and 0.0005 (Fig. 2h). The findings demonstrated that both the HSMF and DFLJ datasets have decreased accuracy. Higher recognition accuracy was achieved in the PSRZ dataset at learning rates of 0.001, 0.0008, and 0.0005. The learning rate of 0.001 outperformed the green tea grade recognition models when the three datasets were compared.

Five learning rate decays StepLR, MultiStepLR, ExponentialLR, CosineAnnealingLR, and ReduceLROnPlateau were compared for their optimization impacts (Fig. 3a–f).

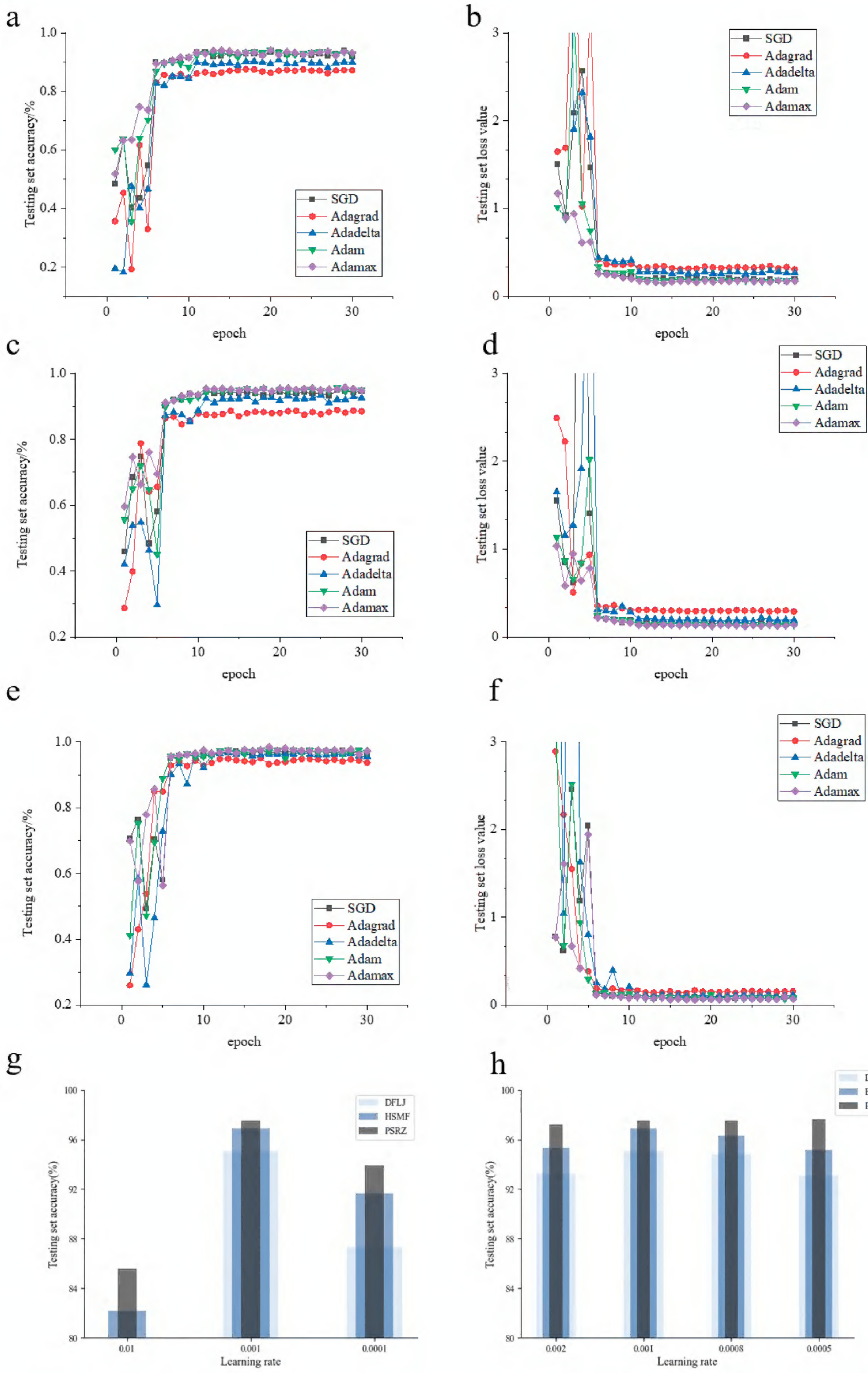

Figure 2. Impact of optimizer and learning rate modification on ResNet18-based green tea grade recognition models. **a.** Accuracy of five optimizers in the DFLJ grade recognition model; **b.** Loss values of five optimizers in the DFLJ model; **c.** Accuracy of five optimizers in the HSMF model; **d.** Loss values of five optimizers in the HSMF model; **e.** Accuracy of five optimizers in the PSRZ model; **f.** Loss values of five optimizers in the PSRZ model; **g.** The impact of initial adjustments to the learning rate in the green grade recognition model; **h.** The impact of additional adjustments to the learning rate in the green tea grade model.

The model recognition accuracy was lower for StepLR and MultiStepLR among the three green teas. The learning rate decay of CosineAnnealingLR had high accuracy in the testing set. However, during training, the accuracy and loss values fluctuated significantly, and the performance lacked sufficient stability. In the DFLJ dataset, ExponentialLR and ReduceLROnPlateau performed better in recognition model, with accuracy of 97.61% and 97.45%, respectively, more than 4% higher than MultiStepLR. In the HSMF dataset, the learning rate decay of ExponentialLR performed better with an accuracy rate of 98.71%, more than 3% higher than MultiStepLR. In the PSRZ dataset, the learning rate decay of exponentialLR and CosineAnnealingLR demonstrated superior recognition accuracy, with all evaluation metrics exceeding 99%. In three green tea grade datasets, the learning rate decay of ExponentialLR performed admirably, demonstrating high accuracy and steady performance during training.

While the learning rate was fixed, the batch size was chosen to optimize network performance. The optimization effects of the five batch sizes of 8, 16, 32, 64, and 128 were compared (Fig. 3g–i). The four performance metrics matched at different batch sizes. The DFLJ dataset performed best with a batch size of 16, followed by a batch size of 32. Both HSMF and PSRZ performed best with a batch size of 32. The precision of HSMF was 98.74%, and the other three metrics were 98.71%. All four performance metrics of PSRZ reached 99.39%. Taken together, batch size 32 was better recognized in all three datasets.

Overfitting optimization

Neural network overfitting was more prevalent when problems like an excess of parameters and insufficient training data were present (Gulzar and Ünal 2025a). To illustrate the impact of overfitting, L2 regularization and Dropout were chosen. With values of 0.01, 0.001, and 0.0001, respectively,

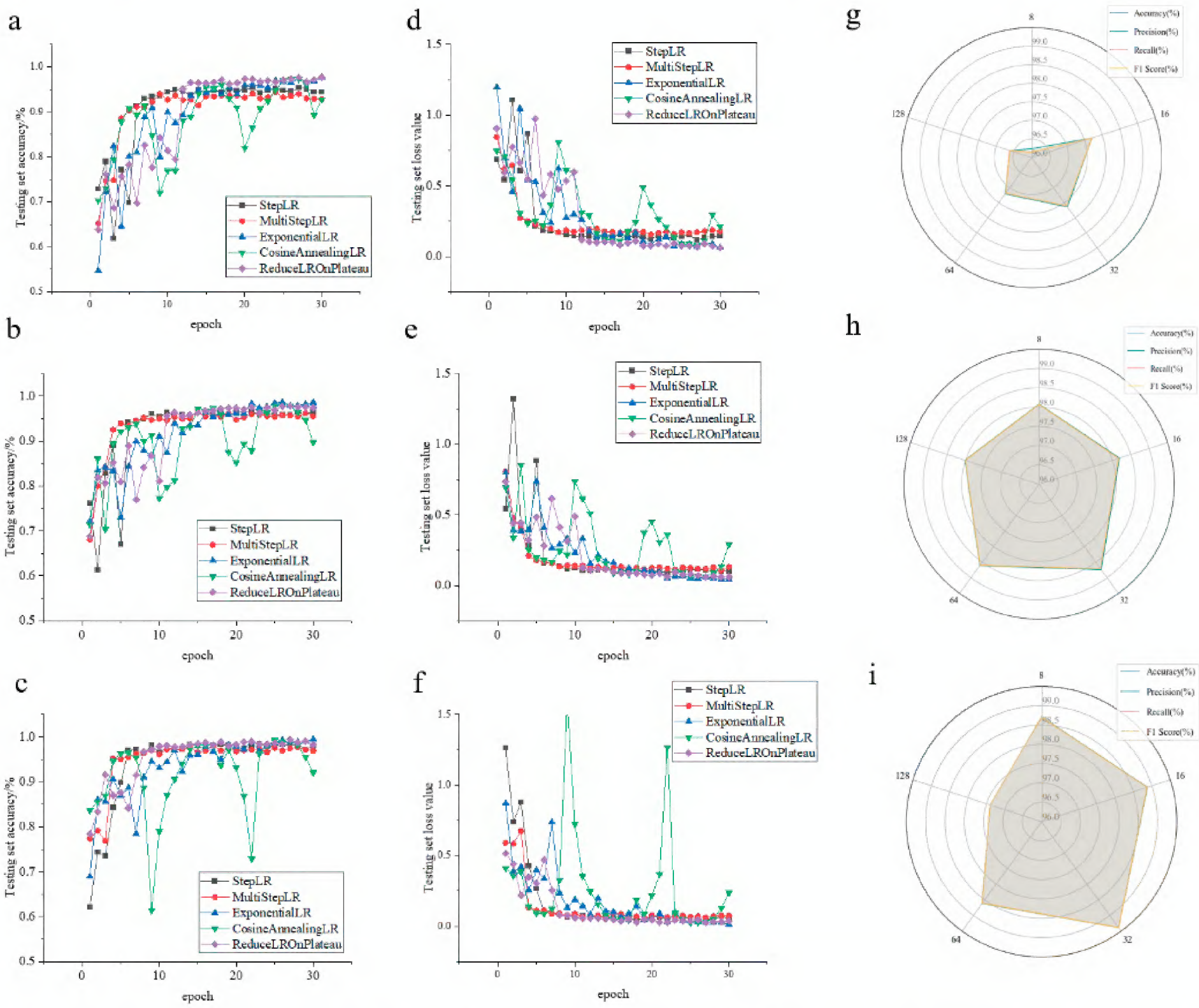

Figure 3. Effects of learning rate decay and batch size adjustment on ResNet18-based green tea grade recognition models. **a.** Accuracy of five learning rate decays in the DFLJ model; **b.** Accuracy of five learning rate decays in the HSMF model; **c.** Accuracy of five learning rate decays in the PSRZ model; **d.** Loss value of five batch sizes in the DFLJ model; **e.** Loss value of five batch sizes in the HSMF model; **f.** Loss value of five batch sizes in the PSRZ model; **g.** Accuracy of five batch sizes in the DFLJ model; **h.** Accuracy of five batch sizes in the HSMF model; **i.** Accuracy of five batch sizes in the PSRZ model.

the L2 regularization was represented by 0.01L, 0.001L, and 0.0001L. 0.1D, 0.3D, and 0.5D represented the dropout regularization parameters of 0.1, 0.3, and 0.5, respectively. The result of adding regularization was compared to the result of not adding it. The 0.01L for the three green tea datasets showed much lower accuracy than the model without regularization, along with slower convergence, higher loss values, and severe oscillations. The accuracy, precision, recall, and F1-score differences of the remaining models were all within 1% of each other. The accuracy improvement for the DFLJ dataset was around 0.5% for 0.0001L and 0.3D. However, the accuracy of the HSMF and PSRZ datasets showed a sharp decline due to L2 regularization and dropout for various parameters. Therefore, in this study, grade recognition without adding regularization was better.

Attention mechanisms significantly increase the accuracy and efficiency of the network by helping it focus on the key elements of the image (Niu et al. 2021). The results of ResNet18 grade recognition models were compared by adding four attentional mechanisms, SE, CBAM, CA, and ECA, and by not adding attentional mechanisms. With the addition of the SE and ECA attention mechanisms, the model recognizes better than without the addition of the attention mechanism. In contrast, adding the ECA attention mechanism improved the performance metrics more significantly.

Evaluation of modeling effects

The ResNet18 model used the AdamX optimizer, 0.001 learning rate, exponential decay strategy, and batch size of 32 for green tea grade recognition with the addition of an ECA attention mechanism. The models recognized correct samples of 1888, 1845, and 1150 in the DFLJ, HSMF, and PSRZ datasets, respectively. The testing set accuracy was 98.28%, 99.19%, and 99.74%, indicating good performance. As shown in the confusion matrix diagram (Fig. 4a–c), the DFLJ dataset featured more incorrectly identified samples in the medium and low grades. Five samples from 2 were incorrectly classed as 3, and two samples were misclassified as 1 and 5, respectively. Four samples in 4 were incorrectly classed as 3, while four samples were incorrectly classified as 5. Seven samples were mistakenly categorized as 4 in 5. All S1 and 3 samples in the HSMF dataset were correctly identified. Three S2 samples were incorrectly classified as S1, and two S2 samples were misclassified as S3. It was incorrect to identify two samples from 1 as S3 and 3, respectively. And two samples were mistakenly assigned to 2. Two samples from 2 were incorrectly assigned to 1, and one sample was incorrectly identified as 3. In the PSRZ dataset, all 3 and 4 samples were accurately identified.

MobileNetV3 model building and green tea grade recognition

Hyperparameter tuning

SGD, Adagrad, Adadelta, Adam, and AdamX optimizers were examined for identification performance of three MobileNetV3-based green tea grade recognition models (Fig. 5). The evaluation metrics indicated that the MobileNetV3 model with the Adagrad optimizer does not

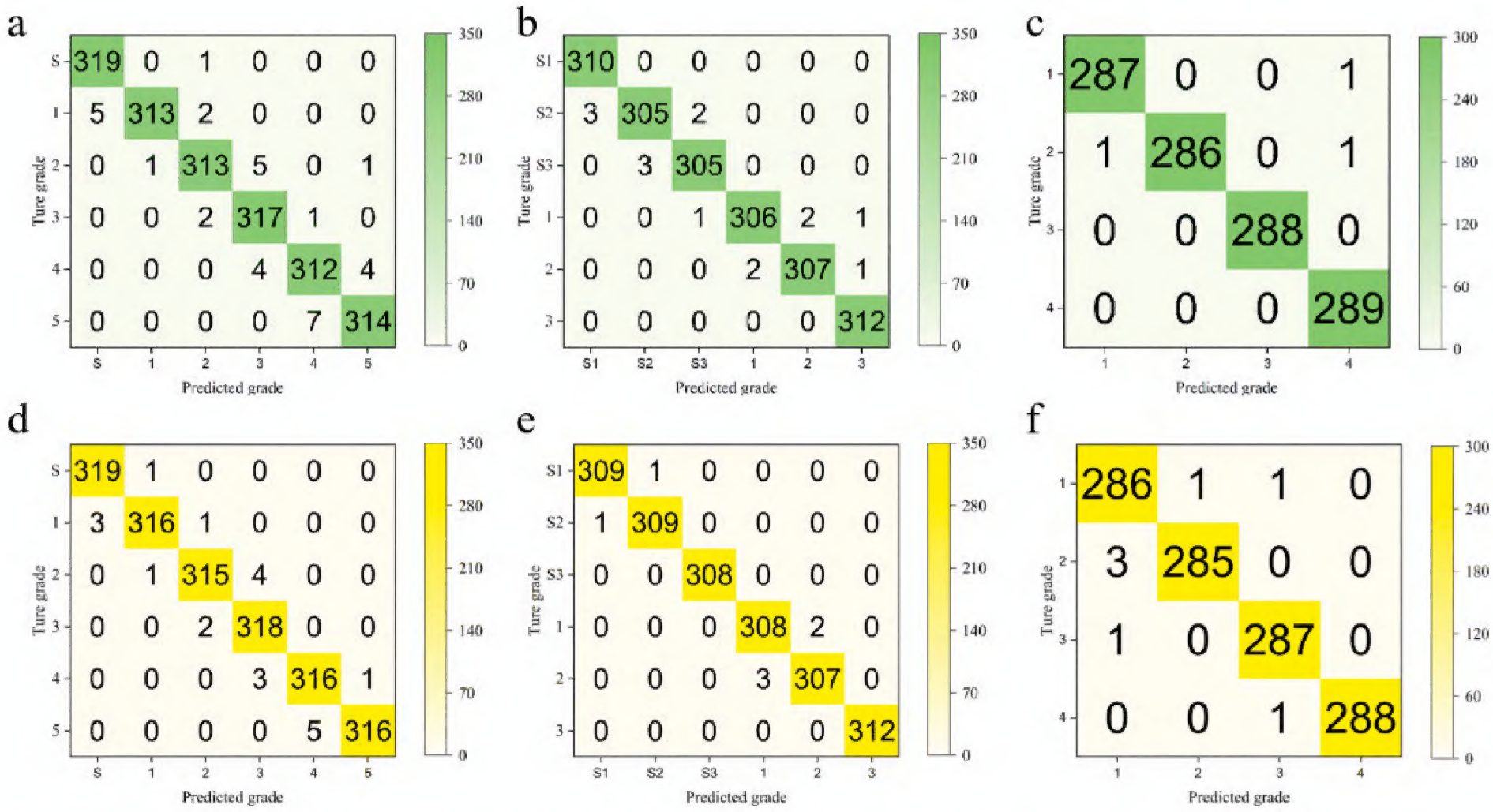

Figure 4. The confusion matrix diagram of the testing set of green tea grade recognition models. **a.** Confusion matrix diagram of ResNet18-based DFLJ model; **b.** Confusion matrix diagram of ResNet18-based HSMF model; **c.** Confusion matrix diagram of ResNet18-based PSRZ model; **d.** Confusion matrix diagram of MobileNetV3-based DFLJ model; **e.** Confusion matrix diagram of MobileNetV3-based HSMF model; **f.** Confusion matrix diagram of MobileNetV3-based PSRZ model.

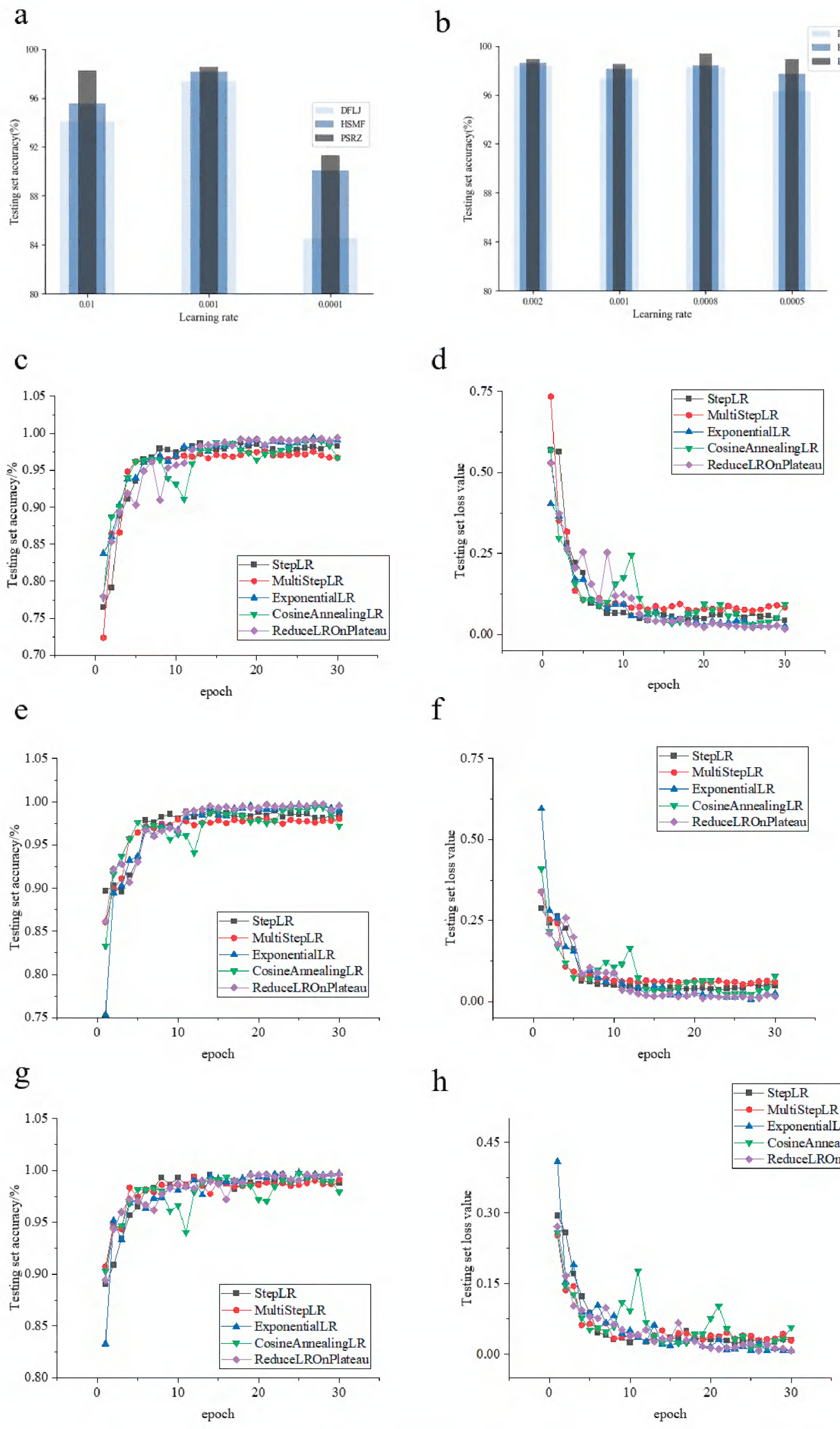

Figure 5. Impact of learning rate and optimizer modification on MobileNetV3-based green tea grade recognition models. **a.** The impact of initial adjustments to the learning rate in the green grade recognition model; **b.** The impact of additional adjustments to the learning rate in the green tea grade model; **c.** Accuracy of five optimizers in the DFLJ grade recognition model; **d.** Loss values of five optimizers in the DFLJ model; **e.** Accuracy of five optimizers in the HSMF model; **f.** Loss values of five optimizers in the HSMF model; **g.** Accuracy of five optimizers in the PSRZ model; **h.** Loss values of five optimizers in the PSRZ model.

perform well. The MobileNetV3 model with the Adam and AdamX optimizers performed better in three green tea datasets, with performance metrics all above 98%.

Three learning rates of 0.01, 0.001, and 0.0001 were established, like in the prior investigation (Fig. 5a). In three datasets of green tea grade, the MobileNetV3 model with learning rates of 0.01 and 0.0001 was less effective. Additionally, the model had higher loss values, worse accuracy, and a delayed rate of convergence. In three green tea datasets, the model using 0.001 as the learning rate performed better in three green tea datasets, with respective accuracy of 97.35%, 98.12%, and 98.53%, respectively. The learning rate was modified to 0.002, 0.0008, and 0.0005 in light of this (Fig. 5b). Compared to a learning rate of 0.001, the accuracy obtained for learning rates of 0.002 and 0.0008 was higher. A learning rate of 0.002 was more accurate than 0.0008 in the DFLJ and HSMF datasets.

Examined and compared the optimization effects of five learning rate decays of StepLR, MultiStepLR, ExponentialLR, CosineAnnealingLR, and ReduceLROnPlateau. The StepLR and MultiStepLR performed less well in distinguishing between the three green tea datasets. In the DFLJ dataset, the learning rate decay of ReduceLROnPlateau was more effective in identification, with an accuracy of 99.53%, which was about 3% higher than MultiStepLR. The learning rate decay of ExponentialLR, CosineAnnealingLR, and ReduceLROnPlateau in the HSMF grade discrimination was better, with accuracy rates above 99%. In the PSRZ dataset, ExponentialLR and CosineAnnealingLR showed a better recognition effect, with evaluation metrics of accuracy, precision, recall, and an F1-score over 99.30%. On the other hand, CosineAnnealingLR displays periodic oscillations with a wider range during training (Fig. 5c–h).

Evaluation of modeling effects

Using the AdamX optimizer, 0.002 learning rate, ExponentialLR learning rate decay, and a batch size of 32, a MobileNetV3-based green tea grade recognition model was built. 1900, 1853, and 1145 samples of DFLJ, HSMF, and PSRZ green tea samples were properly identified, and the testing set recognition accuracies were 98.91%, 99.62%, and 99.39% (Fig. 4d–f). Misclassification in the DFLJ mostly happened in grades 1, 2, 4, and 5. Three samples in 1 were misclassified as S, while one sample was misclassified as 2. One sample in 2 was misclassified as 1, and four samples were misclassified as 3. Three samples for 4 were incorrectly classed as 3, while one sample was incorrectly classified as 5. Five samples of 5 were misclassified as 4. In the HSMF dataset, all the S3 and 3 samples were correctly recognized. S1 and S2 have one misclassification, whereas 1 and 2 have two and three misclassifications, respectively. In the PSRZ dataset, three samples of 2 were misclassified as 1.

Discussion

It has been suggested that the color of needle-shaped green tea combined with machine learning models can effectively evaluate key color indicators of tea appearance. There are many types of green teas in China, and it would be too much work to extract characteristics for each type of tea. Deep learning was introduced in this study without the need to specifically extract tea features of every type. As shown in prior studies, deep learning can significantly improve the accuracy of tea quality recognition model (Ren et al. 2024). To reduce the effect of imaging differences, tea images from different camera devices were collected for analysis to reduce variances brought on by variations in shooting angle, distance, and lighting. Many studies have shown that the ResNet model outperformed the MobileNet model in the recognition task. When it comes to the recognition of Hand Gesture images in various countries, ResNet model was 0.5% more accurate than MobileNet model (Alnuaim et al. 2022). With 98% accuracy, the ResNet model outperforms the MobileNet model in the recognition of images of renal illness (Amanatulla et al. 2024). 143 oolong teas from 5 categories can be classified with 93% accuracy using the RresNet50 model (Zhu et al. 2023). In this study, ResNet18 (embedded with ECA attention) and MobileNetV3 achieved grade recognition rates greater than 98% in all three green teas (Fig. 6). ResNet18-based recognition models of DFLJ, HSMF, and PSRZ had accuracies for the testing sets of 98.28%, 99.19%, and 99.74%, respectively. Based on the MobileNetV3 model, the DFLJ, HSMF, and PSRZ testing sets demonstrated recognition accuracies of 98.91%, 99.62%, and 99.39%, respectively. ResNet18 was slightly better in granular tea recognition, while MobileNetV3 performed better in flat-shaped tea and sparrow-tongue shaped tea. This difference may stem from the fact that MobileNet's depth-separable convolution (Howard et al. 2017) better preserves detailed features such as leaf integrity, vein texture. Naturally, the accuracy of picture identification may be greatly increased by integrating MobileNet with ResNet (Azhar et al. 2024). In the DFLJ, HSMF, and PSRZ datasets, the attentional mechanism increased the ResNet18 model's accuracy by 0.67%, 0.48%, and 0.35%, respectively. By adding an attention mechanism, it was also able to improve the accuracy of machine vision inspection of wheat grains (Zhang et al. 2021). The AdamX optimizer and ExponentialLR learning rate decay strategy were used to demonstrate significant parameter stability, with accuracy fluctuations of <1.5% among different tea forms. Therefore, in the future, these two models are suitable to be deployed in software or specialized equipment used for tea grade identification for tea processing factories, retail stores, and general consumers. On the other hand, the ResNet18 model with the ECA attention mechanism was more than three times larger than the MobileNetV3 model. The MobileNet method was effective in identifying diseases on apple leaves. Just two seconds were needed to recognize a single image. However, the cost of ResNet was nearly four times that of the MobileNet model (Bi et al. 2022). These results demonstrate that lightweight

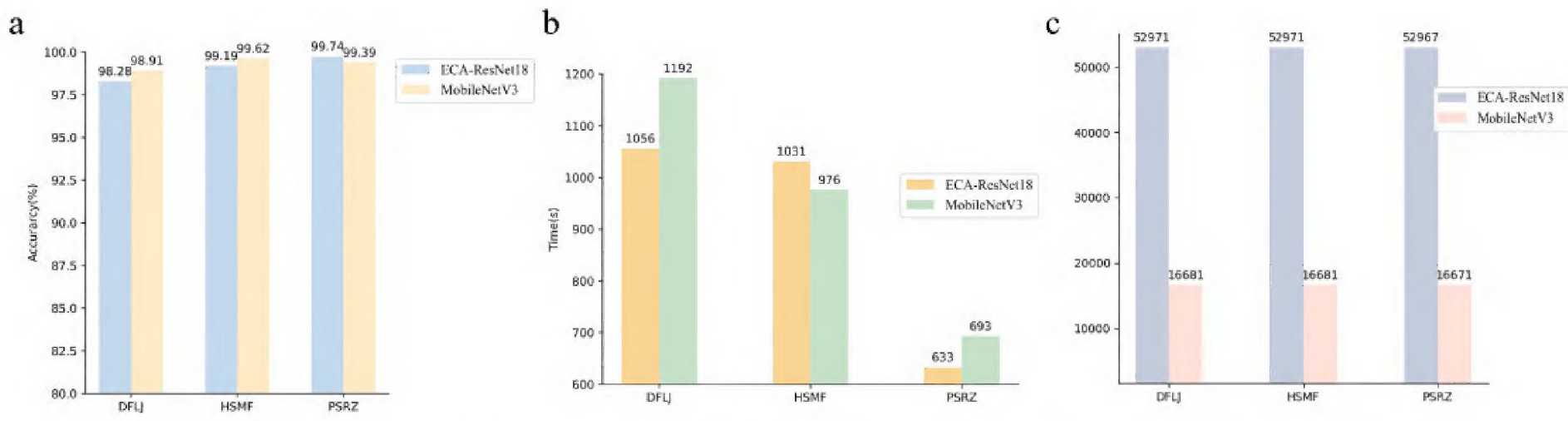

Figure 6. Comparison of green tea grade recognition models based on ResNet18 and MobileNetV3. **a.** The accuracy of the model; **b.** Training time of the model; **c.** Model size.

models can achieve as high an accuracy rate as complex models in the field of agricultural vision tasks. (Mohapatra et al. 2021). However, the current study was limited to grade recognition of known types of green tea. In the future, there is a need to further expand the dataset of green tea images and improve the tea coverage of the recognition model. At the same time, parameter tuning will be optimized in depth to improve the applicability of the model. In addition, we will further explore incorporating tea type identification methods to simultaneously discriminate tea types and grades.

Conclusions

In this study, a deep learning-based computer vision solution was constructed for the problem of recognizing the grades of different appearances of green tea. A dataset containing different grades of flat-shaped DFLJ tea, and sparrow tongue-shaped HSMF tea and granular-shaped PSRZ tea was established to verify the model applicability of deep convolutional networks in tea appearance feature quality recognition. Under the optimal parameters, the ResNet18 model and the MobileNetV3 model were able to stably recognize grades in different shapes of green tea, respectively. The ResNet18-based grade recognition model with AdamX as the optimizer, 0.001 as the learning rate, ExponentialLR as the learning rate decay, and 32 as the batch size, and the inclusion of the ECA attentional mechanism had good performance. The accuracy of the testing set was 98.28% in DFLJ, 99.19% in HSMF, and 99.74% in PSRZ. The MobileNetV3-based model with AdamX as the optimizer, 0.002 as the learning rate, ExponentialLR as the learning rate decay, 32 as the batch size showed good accuracy of grade identification. The accuracy of the testing set was 99.62%, 99.39%, and 98.91%, respectively. The model proposed in this study does not limit the equipment and scenarios of image acquisition, and it does not need to extract the features of tea leaves during the classification process, and it quickly realizes the classification of tea grades.Through comparison, it is found that the ResNet18 model and the MobileNetV3 model have higher accuracies in recognizing the grades of the three shapes of green tea, which is suitable to be generalized to recognize the grades of other green teas. For mobile deployment, the MobileNetV3 model is more suitable for the lightweight model.

Author contributions

Conceptualization, Danyi Huang; methodology, Danyi Huang and Lamei Li ; software, Yimin Xuan and Danyi Huang ; validation, Dongmei Fan; investigation, Dongmei Fan; resources, Dongmei Fan; data curation, Yimin Xuan and Lamei Li ; writing—original draft preparation, Danyi Huang; writing—review and editing, Xiaochang Wang; visualization, Yimin Xuan and Danyi Huang ; supervision, Xiaochang Wang; project administration, Xiaochang Wang; funding acquisition, Xiaochang Wang and Danyi Huang. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the Science and Technology Program of Zhejiang Province (2022C02075); Ph.D. Fund Program of Zhejiang Institute of Economics and Trade(24BSJJ01).

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data collection and analysis

The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgements

The authors would like to thank the editor and reviewers for their assistance and valuable comments.

References

Alnuaim A, Zakariah M, Hatamleh WA, Tarazi H, Amoatey ET (2022) Human-computer interaction with hand gesture recognition using ResNet and MobileNet. Computational Intelligence and Neuroscience 2022(1): 8777355. https://doi.org/10.1155/2022/8777355

Amanatulla M, Swathi G, Pallavi M (2024) MRI scans for Deep Learning-based chronic nephropathy detection: A comparison of CNN, MobileNet, VGG16, and ResNet-50 models. Paper presented at the 2024 5th International Conference for Emerging Technology (INCET). https://doi.org/10.1109/INCET61516.2024.10593144

Amri E, Gulzar Y, Yeafi A, Jendoubi S, Dhawi F, Mir MS (2024) Advancing automatic plant classification system in Saudi Arabia: introducing a novel dataset and ensemble deep learning approach. Modeling Earth Systems and Environment 10(2): 2693–2709. https://doi.org/10.1007/s40808-023-01918-9

Azhar NZB, Teo NHI, Hamzah Rb, Roslan Rb, Maskat R (2024) A Hybrid ResNet — MobileNet Deep Learning Model for Smart Bin Waste Classification. Paper presented at the 2024 5th International Conference on Artificial Intelligence and Data Sciences (AiDAS). https://doi.org/10.1109/AiDAS63860.2024.10730599

Bagherzadeh S, Norouzi MR, Hampa SB, Ghasri A, Kouroshi PT, Hosseininasab S, Zadeh MAG, Nasrabadi AM (2024) A subject-independent portable emotion recognition system using synchrosqueezing wavelet transform maps of EEG signals and ResNet-18. Biomedical Signal Processing and Control 90: 105875. https://doi.org/10.1016/j.bspc.2023.105875

Bakhshipour A, Zareiforoush H, Bagheri I (2020) Application of decision trees and fuzzy inference system for quality classification and modeling of black and green tea based on visual features. Journal of Food Measurement and Characterization 14(3): 1402–1416. https://doi.org/10.1007/s11694-020-00390-8

Bi C, Wang J, Duan Y, Fu B, Kang J-R, Shi Y (2022) MobileNet based apple leaf diseases identification. Mobile Networks and Applications 27: 172–180. https://doi.org/10.1007/s11036-020-01640-1

Chen J, Liu Q, Gao L (2019) Visual tea leaf disease recognition using a convolutional neural network model. Symmetry 11(3): 343. https://doi.org/10.3390/sym11030343

Choi C, Han J, Son Y, Joo S, Kwon S, Y-HL (2023) Green tea extract exhibits antidiabetic effects partly through regulating dipeptidyl peptidase-4 expression in ddipose tissue. The Journal of Nutritional Biochemistry 111: 109173. https://doi.org/10.1016/j.jnutbio.2022.109173

Chouhan SS, Singh UP, Sharma U, Jain S (2024) Classification of different plant species using Deep Learning and Machine Learning algorithms. Wireless Personal Communications 136(4): 2275–2298. https://doi.org/10.1007/s11277-024-11374-y

Connolly K, Batacan RJ, Jackson D, Fenning AS (2024) Effects of Epicatechin on Cardiovascular function in middle-aged diet-induced obese rat models of metabolic syndrome. British Journal of Nutrition 131(4): 593–605. https://doi.org/10.1017/S000711452300209X

Elfatimi E, Eryigit R, Elfatimi L (2021) Beans leaf diseases classification using MobileNet models. IEEE Access 10: 9471–9482. https://doi.org/10.1109/ACCESS.2022.3142817

Gulzar Y (2024) Enhancing soybean classification with modified inception model: A transfer learning approach. Emirates Journal of Food and Agriculture 36: 1–9. https://doi.org/10.3897/ejfa.2024.122928

Gulzar Y, Ünal Z (2025a) Optimizing pear leaf disease detection through PL-DenseNet. Applied Fruit Science 67: 40. https://doi.org/10.1007/s10341-025-01265-2

Gulzar Y, Ünal Z (2025b) Time-sensitive bruise detection in plums using PlmNet with transfer learning. Procedia Computer Science 257: 127–132. https://doi.org/10.1016/j.procs.2025.03.019

Gulzar Y, Ünal Z, Kızıldeniz T, Umar UM (2024) Deep learning-based classification of alfalfa varieties: A comparative study using a custom leaf image dataset. MethodsX 13: 103051. https://doi.org/10.1016/j.mex.2024.103051

Guo J, Liang J, Xia H, Ma C, Lin J, Qiao X (2024) An Improved Inception Network to Classify Black Tea Appearance Quality. Journal of Food Engineering 369: 111931. https://doi.org/10.1016/j.jfoodeng.2023.111931

He Q, Liu Z, Li X, He Y, Lin Z (2024) Detection of the pigment distribution of stacked matcha during processing based on hyperspectral imaging technology. Agriculture 14(11): 2023. https://doi.org/10.3390/agriculture14112033

Hou Z, Jin Y, Gu Z, Zhang R, Su Z, Liu S (2024) (1)H NMR spectroscopy combined with machine-learning algorithm for origin recognition of chinese famous green tea longjing tea. Foods 13(17): 2702. https://doi.org/10.3390/foods13172702

Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv 1704.04861. https://doi.org/10.48550/arXiv.1704.04861

Jiang M, Chen Z (2021) Symmetry detection algorithm to classify the tea grades using artificial intelligence. Microprocessors and Microsystems 81: 103738. https://doi.org/10.1016/j.micpro.2020.103738

Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet Classification with Deep Convolutional Nural Networks.

Li L, Xie S, Ning J, Chen Q, Zhang Z (2018) Evaluating green tea quality based on multisensor data fusion combining hyperspectral imaging and olfactory visualization systems. Science of Food and Agriculture 99(4): 1787–1794. https://doi.org/10.1002/jsfa.9371

Li L, Chen Y, Dong S, Shen J, Cao S, Cui Q, Song Y, Ning J (2023) Rapid and comprehensive grade evaluation of keemun black tea using efficient multidimensional data fusion. Food Chemistry X 20: 100924. https://doi.org/10.1016/j.fochx.2023.100924

Liu Z, Yu W, Chen W (2010) Short Text Feature Selection for Micro-blog Mining. Paper presented at the 2010 International Conference on Computational Intelligence and Software Engineering. https://doi.org/10.1109/CISE.2010.5677015

Lu L, Wang L, Liu R, Zhang Y, Zheng X, Lu J, Wang X, Ye J (2024) An efficient artificial intelligence algorithm for predicting the sensory quality of green and black teas based on the key chemical indices. Food Chemistry 441: 138341. https://doi.org/10.1016/j.foodchem.2023.138341

Madhura MG, Shweta RD, Kumar BV, Savithri D, Gajalakshmi S (2016) Antifungal effect of green tea extracts on oral candida species: An *in vitro* study. Advanced Clinical and Research Insights 3: 1–4. https://doi.org/10.15713/ins.jcri.93

Mohapatra S, Abhishek NVS, Bardhan D, Ghosh AA, Mohanty S (2021) Comparison of MobileNet and ResNet CNN Architectures in the CNN-Based Skin Cancer Classifier Model. In: Mohanty SN, Nalinipriya G, Jena OP, Sarkar A (Eds) Machine Learning for Healthcare Applications, Vol. 11: 169–186. https://doi.org/10.1002/9781119792611.ch11

Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452: 48–62. https://doi.org/10.1016/j.neucom.2021.03.091

Ren G, Wu R, Yin L, Zhang Z, Ning J (2024) Description of tea quality using deep learning and multi-sensor feature fusion. Journal of Food Composition and Analysis 126: 105924. https://doi.org/10.1016/j.jfca.2023.105924

Santos MdFaC, Rech KS, Dutra LM, Menezes LRA, Santos ADdC, Nagata N, Stefanello MEA, Barison A (2023) (1)H HR-MAS NMR chemical profile and chemometric analysis as a tool for quality control of different cultivars of green tea (*Camellia sinensis*). Food Chemstry 408: 135016. https://doi.org/10.1016/j.foodchem.2022.135016

Tang Z, Li M, Wang X (2020) Mapping tea plantations from VHR images using OBIA and convolutional neural networks. Remote Sensing 12(18): 2935. https://doi.org/10.3390/rs12182935

Wang S-M, Yu C-P, Ma J-H, Ouyang J-X, Zhao Z-M, Xuan Y-M, Fan D-M, Yu J-F, Wang X-C, Zheng X-Q (2024a) Tea yield estimation using UAV images and deep learning. Industrial Crops & Products 212: 118358. https://doi.org/10.1016/j.indcrop.2024.118358

Wang Y, Ren Y, Kang S, Yin C, Shi Y, Men H (2024b) Identification of tea quality at different picking periods: a hyperspectral system coupled with a multibranch kernel attention network. Food Chemstry 433: 137307. https://doi.org/10.1016/j.foodchem.2023.137307

Wu D, Yang H, Chen X, He Y, Li X (2008) Application of image texture for the sorting of tea categories using multi-spectral imaging technique and support vector machine. Food Engineering 88(4): 474–483. https://doi.org/10.1016/j.jfoodeng.2008.03.005

Wu L, Liu H, Ye C, Wu Y (2024) Development of a premium tea-picking robot incorporating deep learning and computer vision for leaf detection. Applied Sciences 14(13): 5748. https://doi.org/10.3390/app14135748

Yang B, Jiang J, Zhang H, Han Z, Lei X, Chen X, Xiao Y, Ndombi SN, Zhu X, Fang W (2023) Tea quality estimation based on multi-source information from leaf and soil using machine learning algorithm. Food Chemistry: X 20: 100975. https://doi.org/10.1016/j.fochx.2023.100975

Yang CS, Jin H, Guan F, Chen Y-K, Wang H (2020) Cancer preventive activities of tea polyphenols. Journal of Food and Drug Analysis 20(1): 37. https://doi.org/10.38212/2224-6614.2103

Zhang C, Wang J, Yan T, Lu X, Lu G, Tang X, Huang B (2023) An instance-based deep transfer learning method for quality identification of longjing tea from multiple geographical origins. Complex & Intelligent Systems 9(3): 3409–3428. https://doi.org/10.1007/s40747-023-01024-4

Zhang W, Ma H, Li X, Liu X, Jiao J, Zhang P, Gu L, Wang Q, Bao W, Cao S (2021) Imperfect wheat grain recognition combined with an attention mechanism and residual network. Applied Sciences 11(11): 5139. https://doi.org/10.3390/app11115139

Zhou L, Zhang C, Liu F, Qiu Z, He Y (2019) Application of deep learning in food: A review. Comprehensive Reviews in Food Science and Food Safety 18(6): 1793–1811. https://doi.org/10.1111/1541-4337.12492

Zhu Y, Chen S, Yin H, Han X, Xu M, Wang W, Zhang Y, Feng X, Liu Y (2023) Classification of oolong tea varieties based on computer vision and convolutional neural networks. Journal of the Science of Food and Agriculture 104(3): 1630–1637. https://doi.org/10.1002/jsfa.13049